A probabilistic approach for damage identification and crack mode classification in reinforced concrete structures

Alireza Farhidzadeh, Salvatore Salamone and Puneet Singla

Journal of Intelligent Material Systems and Structures published online 18 April 2013
DOI: 10.1177/1045389X13484101

Original Article

Journal of Intelligent Material Systems and Structures

Journal of Intelligent Material Systems and Structures
0(0) 1–14

DOI: 10.1177/1045389X13484101
jim.sagepub.com
SAGE

A probabilistic approach for damage identification and crack mode classification in reinforced concrete structures

Alireza Farhidzadeh[1], Salvatore Salamone[1] and Puneet Singla[2]

Abstract

Reinforced concrete is subjected to deterioration due to aging, increased load, and natural hazards. To minimize the maintenance costs and to increase the operation lifetime, researchers and practitioners are increasingly interested in improving current nondestructive evaluation technologies or building advanced structural health monitoring strategies. Acoustic emission methods offer an attractive solution for nondestructive evaluation/structural health monitoring of reinforced concrete structures. In particular, monitoring the development of cracks is of large interest because their properties reflect not only the condition of concrete as material but also the condition of the entire system at structural level. This article presents a new probabilistic approach based on Gaussian mixture modeling of acoustic emission to classify crack modes in reinforced concrete structures. Experimental results obtained in a full-scale reinforced concrete shear wall subjected to reversed cyclic loading are used to demonstrate and validate the proposed approach.

Keywords

Acoustic emission, damage identification, Gaussian mixture modeling, crack classification, reinforced concrete, cluster analysis

Introduction

Over the past century, reinforced concrete (RC) has been widely used in civil structures such as buildings, bridges, nuclear power plants, or dams whose performance and function are vital to the security of our society. Unfortunately, many of these structures are facing an increasing number of challenges such as aging or natural events (i.e. earthquakes, hurricanes, and tsunami) that can jeopardize their safety and serviceability. To secure the overall soundness of these structures, proper assessment is crucial. In particular, cracks are of large interest in engineering practice (Jahanshahi et al., 2013), since their properties reflect not only the condition of concrete as a material but also the condition of the entire system at the structural level. Methodologies based on acoustic emission (AE) offer an attractive solution for monitoring the nucleation and growth of cracks in RC structures. AEs are stress waves caused by sudden strain releases, such as internal fractures or concrete cracking (Grosse and Ohtsu, 2008). In general, a sparse array of piezoelectric transducers is placed in contact with the structure to detect these waves while some features are extracted from the detected AE signals to assess its structural integrity. Extracting features from AE signals is usually referred to as parameter-based technique. A threshold is used as a reference for the AE features. In general, this threshold is set somewhat above the background noise. Some of the most common features include peak amplitude, rise time, duration, and count (see Figure 1). The AE peak amplitude is defined as the maximum amplitude of the signal. The number of times the signal rises and crosses the threshold is the count of the AE event. The time period between the rising edge of the first count and the

[1]Smart Structures Research Laboratory (SSRL), Department of Civil, Structural, and Environmental Engineering, University at Buffalo, The State University of New York, Buffalo, NY, USA
[2]Lab for Autonomous and Intelligent Robotic Systems (LAIRS), Department of Mechanical and Aerospace Engineering, University at Buffalo, The State University of New York, Buffalo, NY, USA

Corresponding author:
Salvatore Salamone, Smart Structures Research Laboratory (SSRL), Department of Civil, Structural, and Environmental Engineering, University at Buffalo, The State University of New York, 212 Ketter Hall, Buffalo, NY 14260, USA.
Email: ssalamon@buffalo.edu

falling edge of the last count is the duration of the AE event. Finally, the time period between the rising edge of the first count and the peak of the AE event is defined to be the rise time (Miller et al., 2005). The past few years have seen a surge of research aimed at determining functional or correlation interdependencies between crack characteristics and AE features (Carpinteri et al., 2007; Farhidzadeh et al., 2012a; Ohtsu, 2010; Salamone et al., 2011, 2012; Vanniamparambil et al., 2012), including *b*-value analysis (Farhidzadeh et al., 2012b, 2013; Kurz et al., 2006; Shiotani et al., 2001), source localization (Dehghan-Niri et al., 2013; Grosse et al., 1997), and AE energy (Benavent-Climent et al., 2011; Colombo et al., 2005).

In this article, a new approach based on a Gaussian mixture modeling (GMM) is presented to perform crack mode classification in RC structures. GMM is an unsupervised distribution–based classification technique (Bilmes, 1998) that has been successfully used in many fields, including sound recognition (Reynolds et al., 2000; Reynolds and Rose, 1995), image processing (Permuter et al., 2006), dynamical systems and tracking (Povinelli et al., 2004; Stauffer and Grimson, 1999; Terejanu et al., 2008), and text recognition (Richiardi and Drygajlo, 2003); however, the use of this technique for AE-based structural health monitoring (SHM) has not been investigated.

This article is organized as follows: in the next section, we briefly discuss the crack classification method developed by the Japan Construction and Material Standard (JCMS) (JCMS-IIIB5706, 2003) for RC structures, and its shortcomings are highlighted. The theoretical aspects of the GMM algorithm are then described. The following sections describe the experimental study used to validate the proposed algorithm along with the results. At the end, conclusions are summarized.

Conventional crack mode classification

In the last years, several efforts have been made to correlate AE features with the cracking mode in RC members (Aggelis, 2011; Beck et al., 2003; ElBatanouny et al., 2012; Farhidzadeh et al., 2012a; Ohno and Ohtsu, 2010; Shigeishi, 2001). Cracks can adversely affect structural performance in various ways such as durability and serviceability. In general, in loading such structures until failure, tensile cracks as a result of pure tensile stress develop at the initial stages of loading, while shear cracks dominate later (Yuyama et al., 1999). Therefore, it may be beneficial to monitor the mode of cracks, as it may lead to a better prediction of the structural performance and eventually provide an early warning for the planning and implementation of remedial actions to the structure at a point where it is less expensive and invasive than when its structural

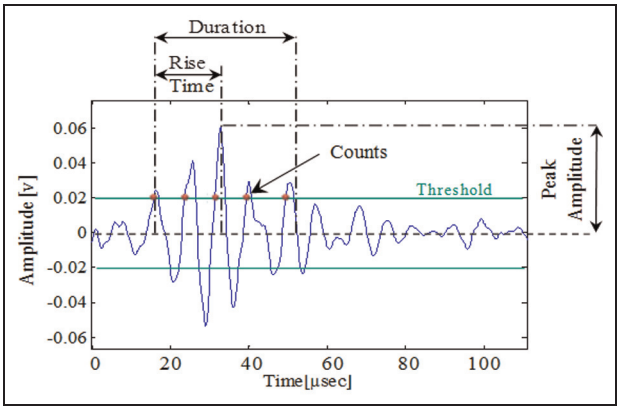

Figure 1. AE parameters in an AE signal.
AE: acoustic emission.

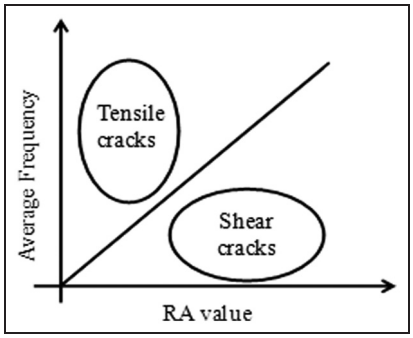

Figure 2. Conventional crack classification in JCMS-IIIB5706 code.

performance has been seriously compromised. A methodology for monitoring the crack propagation in RC structures has been proposed by the JCMS. This methodology is based on two AE parameters, namely, "average frequency (AF)" and "RA" value, which are defined as (see Figure 1)

$$\mathrm{RA} = (\text{Rise time})/(\text{Peak amplitude}) \quad (1)$$

$$\mathrm{AF} = (\text{Counts})/\text{Duration} \quad (2)$$

According to the JCMS, AE sources can be classified into tensile cracks and shear cracks, based on the relationship between RA values and average frequencies, as shown in Figure 2. However, a defined criterion on the proportion of these two parameters has not been confirmed yet (Ohno and Ohtsu, 2010). In fact, AE measurements are mostly random data and, in general, nonlinearly separable. Therefore, the distribution characteristics of the data set should be considered for developing a more robust classification algorithm. In this article, a new probabilistic approach based on a GMM is proposed to take into account data distribution properties and classify AE sources into two dominant clusters, namely, shear and tensile.

Gaussian mixture modeling

A GMM is a parametric probability density function represented as a weighted sum of M component Gaussian densities. For a D-dimensional measurement, training, or feature vector $\vec{x}$, the mixture density is defined as (Reynolds and Rose, 1995)

$$p(\vec{x}|\lambda) = \sum_{i=1}^{M} \omega_i \mathcal{N}_i(\vec{x}|\vec{\mu}_i, \sum_i) \quad i = 1, \ldots, M \tag{3}$$

where ω_i are the mixture weights and $\mathcal{N}_i(\vec{x}|\vec{\mu}_i, \sum_i)$ are the unimodal component Gaussian (Normal) densities. Each component density is a D-variate Gaussian function of the form:

$$\mathcal{N}_i(\vec{x}|\vec{\mu}_i, \sum_i) = \frac{1}{(2\pi)^{D/2}\left|\sum_i\right|^{1/2}} \exp\left\{-\frac{1}{2}(\vec{x}-\vec{\mu}_i)^T\left(\sum_i\right)^{-1}(\vec{x}-\vec{\mu}_i)\right\} \tag{4}$$

with $D \times 1$ mean vector $\vec{\mu}_i$ and $D \times D$ covariance matrix $\sum_i$. The mixture weights satisfy the constraint that $\sum_{i=1}^{M} \omega_i = 1$. The complete Gaussian mixture model is parameterized by the mean vectors, covariance matrices, and mixture weights from all component densities. These parameters are together represented as

$$\lambda = \{\omega_i, \vec{\mu}_i, \sum_i\}, \quad i = 1, \ldots, M \tag{5}$$

For a GMM-based classification system, the goal of the model training is to estimate the parameters of the GMM, λ, so that the Gaussian mixture density can best match the distribution of the feature vector $\vec{x}$. The next step is to find the best estimate for λ.

Maximum likelihood (ML) estimation is one of the most popular and well-established methods available for estimating ω_i, $\vec{\mu}_i$, and $\sum_i$. The aim of ML estimation is to find the model parameters, which maximize the likelihood of the GMM, given the training data. For a sequence of T training vectors $X = \{\vec{x}_1, \ldots, \vec{x}_T\}$, the GMM likelihood, assuming independence between the vectors, can be written as

$$p(X|\lambda) = \prod_{t=1}^{T} p(\vec{x}_t|\lambda) \tag{6}$$

Due to nonlinearity of this expression as a function of λ, direct maximization (i.e. setting first derivative equal to zero and constraining the second derivative to be positive) is not computationally tractable. Instead ML parameters can be obtained iteratively by the expectation maximization (EM) algorithm (Dempster et al., 1977).

The EM algorithm is an iterative procedure that maximizes the likelihood generated by each GMM (Hastie et al., 2009). The basic idea of the EM algorithm is beginning with an initial model λ^k, then estimate a new model λ^{k+1}, such that $p(X|\lambda^{k+1}) > p(X|\lambda^k)$ (Reynolds and Rose, 1995). The new model then becomes the initial model for the next iteration, and the process is repeated until some convergence threshold is reached (i.e. 10^{-6} for log-likelihood). This algorithm consists of two steps, expectation and maximization, which guarantee a monotonic increase in the likelihood value of the model (Dempster et al., 1977). The expectation step results in the a posteriori probability for component i, which is defined as the probability that the state is i and that the mth Gaussian mixture accounts for the observation $\vec{x}_t$, given the kth re-estimated model λ^k

$$\Pr(i|\vec{x}_t, \lambda^k) = \frac{\omega_i \mathcal{N}(\vec{x}_t|\vec{\mu}_i^k, \Sigma_i^k)}{\sum_{j=1}^{M} \omega_j \mathcal{N}(\vec{x}_t|\vec{\mu}_j^k, \Sigma_j^k)} \tag{7}$$

The maximization step returns the distribution parameter with these components (Reynolds and Rose, 1995)

$$\omega_i^{k+1} = \frac{1}{T}\sum_{t=1}^{T} \Pr(i|\vec{x}_t, \lambda^k) \tag{8}$$

$$\vec{\mu}_i^{k+1} = \frac{\left(\sum_{t=1}^{T} \Pr(i|\vec{x}_t, \lambda^k)\vec{x}_t\right)}{\sum_{t=1}^{T} \Pr(i|\vec{x}_t, \lambda^k)} \tag{9}$$

$$\sum_i^{k+1} = \frac{\left(\sum_{t=1}^{T} \Pr(i|\vec{x}_t, \lambda^k) \times (\vec{x}_t - \vec{\mu}_i^{k+1})(\vec{x}_t - \vec{\mu}_i^{k+1})^T\right)}{\sum_{t=1}^{T} \Pr(i|\vec{x}_t, \lambda^k)} \tag{10}$$

The use of a GMM may be motivated by the intuitive notion that the individual component densities may model some underlying set of hidden classes (Rabiner and Juang, 1986; Rammohan and Taha, 2005; Reynolds and Rose, 1995). This possibility motivates application of GMM for crack mode classification in concrete structures with two hidden classes, that is, shear and tensile crack (i.e. $M = 2$) in a two-dimensional measurement vector, $\vec{x} = (RA, AF)$.

To implement a system for crack mode classification in concrete structures, the feature (or measurement) is a 2D vector (i.e. $\vec{x} = (RA, AF)$), the sequence of T training vectors is $X = \{\vec{x}_1 = (RA_1, AF_1), \ldots, \vec{x}_t = (RA_t, AF_t), \ldots, \vec{x}_T = (RA_T, AF_T)\}$, and the latent or hidden classes are $I = \{1,2\}$ for tensile and shear mode, respectively. We aim at estimating the parameters of the GMM (weight, mean, and covariance matrix for each hidden class), which best matches the distribution of the

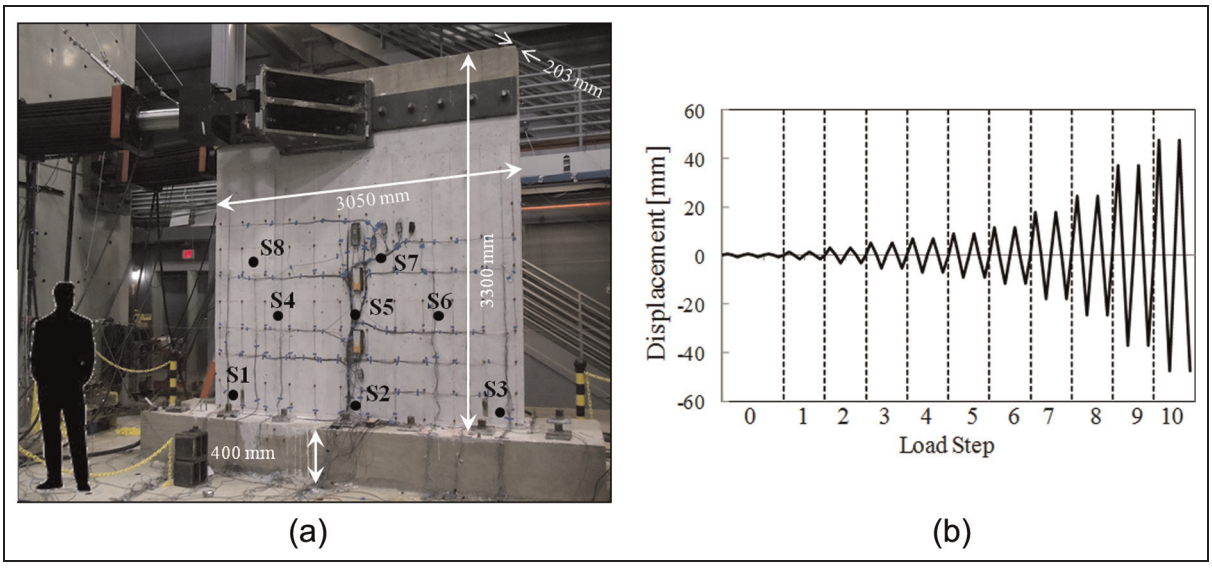

Figure 3. Experimental test: (a) test setup view, AE sensor layout (S1–S8), and dimensions; (b) load protocol.
AE: acoustic emission.

training feature vectors. The overall procedure can be summarized as

1. Initializing the parameters in λ. The initial guess of the parameters of the state-dependent Gaussian mixtures is obtained using vector quantization with two codes (Zhou et al., 2009).
2. Applying equation (7) to get $\Pr(i|x_t, \lambda^k)$.
3. Using $\Pr(i|x_t, \lambda^k)$ to compute a better estimate for the parameter λ^{k+1} (i.e. applying equations (8) to (10)).
4. Iterating steps 2 and 3 until convergence.

Experimental study

The test specimen was a large-scale rectangular RC shear wall with a height-to-width ratio of 0.94, designed based on ACI318-08, Chapter 21, earthquake-resistance structures (ACI Committee 318, 2008). The width, thickness, and effective height (from foundation to the centerline of loading) were 305 (120 in), 20.3 (8 in), and 287 cm (112.8 in), respectively (Luna, 2013; Rocks, 2012). The horizontal and vertical reinforcement ratios were 0.67%. Compressive strength of concrete on the day of test was 24.8 MPa, and the yield and ultimate strength of the reinforcing bars were 464 and 708 MPa, respectively (Rocks, 2012). The main components of the AE system included an eight-channel high-speed data acquisition board (Physical Acoustics Corporation Micro-II PAC) and a dedicated software for signal processing and storage (AEwin). The test specimen was instrumented with eight R15α sensors with resonance frequency of 150 kHz and operating frequency range of 50–400 kHz. They were attached to one face of the wall using hot glue. Preamplifiers were set at 40 dB gain, analog bandpass filters were adjusted in the interval of 20–400 kHz, and the threshold level was set to 35 dB. The experimental setup and load protocol are shown in Figure 3. The lateral load was applied to the specimen by two horizontally inclined high force capacity actuators. The specimen was subjected to a displacement controlled reversed cyclic loading with a loading rate of 0.6 mm/s. The loading protocol consisted of 11 load steps (LSs). In particular, the first load step (LS0) had three cycles, whereas the subsequent load steps had two cycles. LS0 was carried out to verify the functionality of the experimental setup. The crack classification analysis was carried out using data recorded between LS2 and LS9. In fact, a very low AE activity was recorded prior to LS2, whereas in LS10 the specimen was severely damaged resulting in a significant amount of AE activity mostly due to concrete spalling, wall-foundation sliding, and rebar yielding. The force–displacement hysteresis loops and the corresponding backbone curves are illustrated in Figure 4. An initial lateral stiffness reduction of the specimen, caused by the nucleation of microcracks and hairline cracks (Farhidzadeh et al., 2012a, Rocks, 2012), was observed after LS2. In a previous work, authors demonstrated, using Sifted *b*-value (S*b*) analysis (Farhidzadeh et al., 2012a, Farhidzadeh and Salamone, 2012), that a microcrack–macrocrack transition occurred in LS3. Then the wall behaved linearly up to LS6. Significant nonlinear response occurred in LS7, and finally, the wall reached its ultimate strength at LS9. Figure 5 exhibits cracking at peak deformation of the first cycle in load steps 2, 5, and 9 to manifest the development of orientation and positions of cracks. The cracking patterns at the negative load peaks were consistent with these pictures due to the symmetric behavior of the wall. Tensile (horizontal)

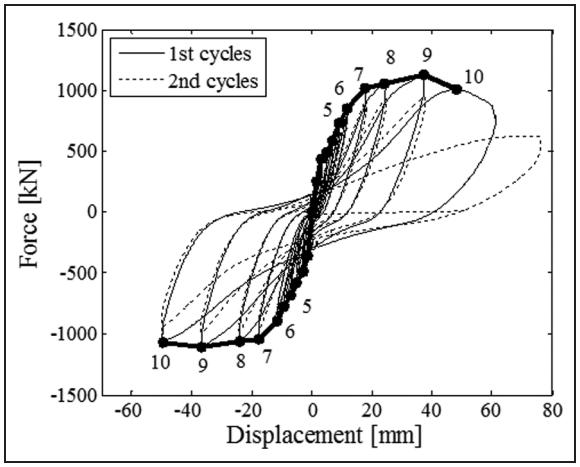

Figure 4. Force–displacement hysteresis loops and backbone curves.

and shear cracks (diagonal) are mapped on the photography by solid and dash lines, respectively. In the initial load steps, the cracks started from tensile mode and then they were followed by diagonal shear cracks in the intermediate load steps. The diagonal cracks became dominant at the final load steps (see LS9 in Figure 5). In addition, maximum and modal (i.e. appearing most frequently) crack widths are marked in each picture with a square dot (□) and with a circle (○), respectively.

Experimental results

In this section, the results of the crack classification using the proposed GMM algorithm are presented. For each load step, the two features defined in equations (1) and (2) (i.e. RA and AF) were extracted from each AE signal detected. Then, a moving average with span and lag of 70 hits was carried out to reduce scatter (Aggelis et al., 2011; Ohtsu, 2010; Soulioti et al., 2009). As a result, the data set $\boldsymbol{X}$, consisting of data points $x_i = (RA_i, AF_i)$, was generated, that is

$$X = \{(RA_1, AF_1), (RA_2, AF_2), \ \ldots, (RA_T, AF_T), \quad x_t \in \mathrm{R}^2\}.$$

In order to apply the GMM algorithm, the number of classes should be defined. The sufficiency of two dominant clusters selected in this work can be investigated using the Bayesian information criterion (BIC), which is

$$\mathrm{BIC} = -2\ln(L) + k \times \ln(T) \tag{11}$$

where L is the maximized value of the likelihood function for the estimated model, T is the number of observations, and k is the number of free parameters to be estimated for each GMM. The BIC criterion provides a tradeoff between model accuracy and complexity of the model. It should be noticed that ideally the data set can be fitted more accurately by increasing the number of mixture components. However, this leads to the estimation of more number of parameters and hence increases model complexity. Furthermore, this can also lead to overfitting the data and experiment noise. BIC can help to select the optimum number of clusters required while maintaining a balance between accuracy and model complexity. Figure 6 illustrates the variation of BIC by additional number of clusters for each load step. It can be observed that from one cluster to two clusters, the BIC decreases significantly and follow a fairly constant trend by including additional clusters. Moreover, in order to explore the efficacy of adding a cluster with a better resolution, the BIC change ratio (ξ_{M+1}) is calculated according to the following equation when M component Gaussian densities increase to $M + 1$

$$\xi_{M+1} = \frac{\mathrm{BIC}_M - \mathrm{BIC}_{M+1}}{\mathrm{BIC}_M} \tag{12}$$

where BIC_M is the BIC with M mixture components. Figure 7 depicts the variation of ξ versus additional

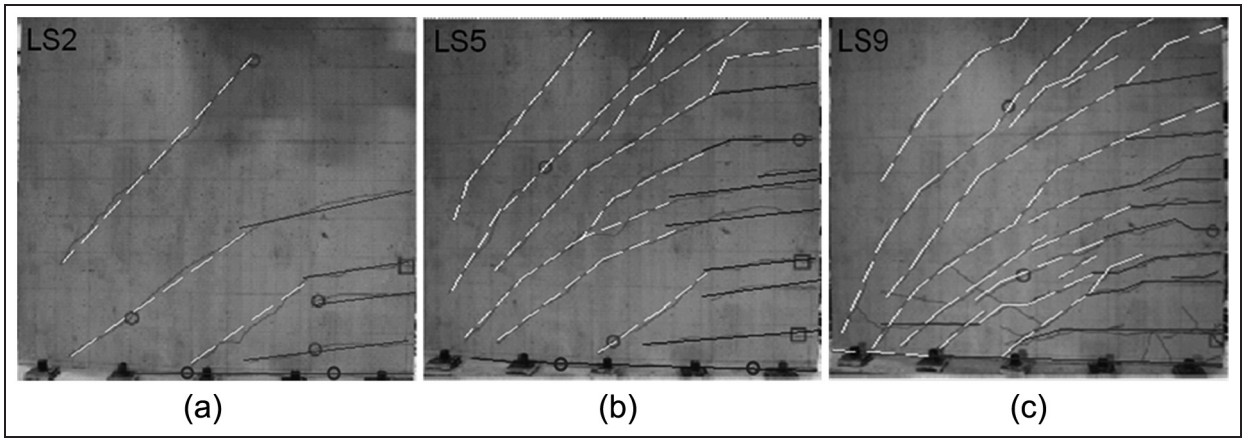

Figure 5. Cracks position, orientation, and mode in the positive peak of first cycles of load steps: (a) LS2, (b) LS5 and (c) LS9.

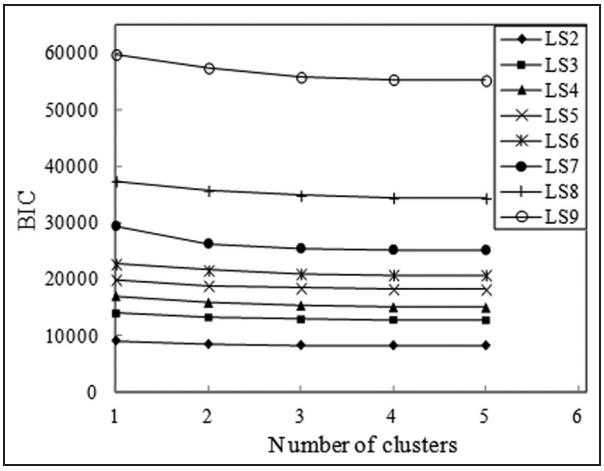

Figure 6. BIC for each load step.
BIC: Bayesian information criterion.

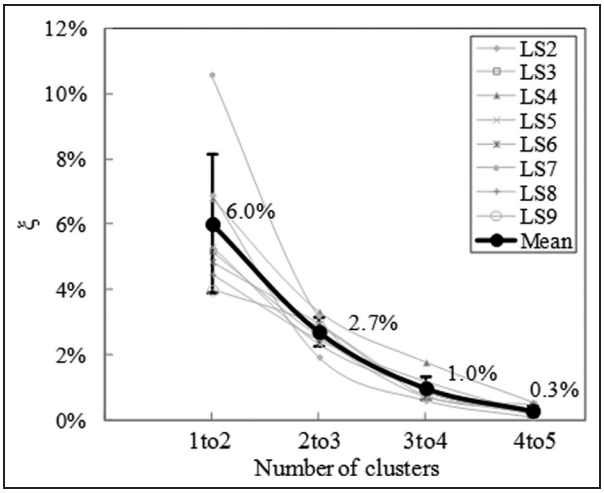

Figure 7. Percentage of reduction in BIC by adding one cluster to the previous number of clusters. The bars represent the standard deviation.
BIC: Bayesian information criterion.

mixture component. The BIC reduces from one cluster to two clusters ranging from 4% to 11% depending on the load step with an average of 6%. However, this reduction is not significant in any load step by adding another mixture component because the improvement is less than 3%. Therefore, considering the computation expenses, two clusters can efficiently fit the data distribution. This outcome is consistent with other studies in which the selection of two clusters has been suggested (Aggelis, 2011; JCMS-IIIB5706, 2003; Ohtsu, 2010; Ohno and Ohtsu, 2010).

To compare the outcome of GMM with conventional JCMS approach, Figure 8 illustrates the feature vector $\boldsymbol{X}$ recorded during the LS2 as an example. As is said earlier, the conventional classification scheme (Figure 8(a)) does not provide a well-established criterion to define the separator line, here indicated by dashed lines; therefore, this method cannot provide conclusive results. On the other hand, Figure 8(b) displays the contour plot of the Gaussian mixture density function. The following observations can be made from this figure: (a) despite Figure 2 that implies two mutually exclusive categories of shear and tensile, it can be observed that these classes are intersected and have a mixed Gaussian density function (with a larger weight for the tensile class in this load step) and (b) although a straight line may not be an appropriate

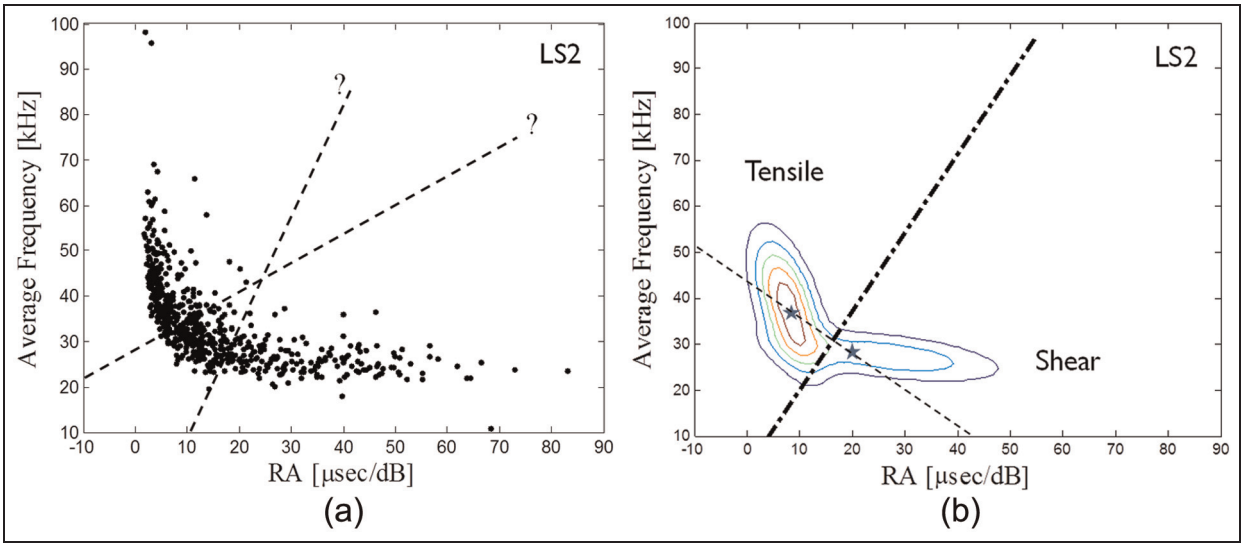

Figure 8. Average frequency versus RA at LS2: (a) ambiguity about separator line based on JCMS code and (b) GMM result and the newly defined separator line.
JCMS: Japan Construction and Material Standard; GMM: Gaussian mixture modeling.

boundary for classification, the GMM can smooth the path toward identifying the optimum discriminator line, which is simply the segment bisector of a virtual line connecting the means of the two classes (dash-dot line). In this article, however, a new criterion will be later proposed instead of a straight line to discriminate the two classes.

To illustrate the effect of the window span size on the GMM results, Figure 9 shows the feature vector $\boldsymbol{X}$ in LS2 obtained using different window span sizes (i.e. $n = 30, 50, 70, 100$) and its corresponding GMM results. It can be observed that, although the span of the moving window reduces the number of data (Figure 9(a), (c), (e), and (g)), the shape of the GMM results and the estimated mixture weights ω_i are not significantly affected.

To monitor the overall behavior of the wall during the experiment, the results of the GMM for the load steps 2, 3, 5, 6, 7, and 9 are illustrated in Figure 10. These results were obtained using data from all eight sensors. The red regions represent higher probability density all the way through the blue regions, which depict smaller densities. The tensile and shear classes are encircled by dashed ellipses representing the confidence interval of 0.9. By definition, confidence interval gives an estimated range of values, which is likely to include an unknown population parameter (Easton and McColl, 1997) and vary between 0 and 1. Three stages can be identified in Figure 10: (a) the dominance of tensile cracks during the initial load steps (i.e. LS2 and LS3). In these load steps, the concentration of feature vectors is around the mean of the tensile class. (b) A transition stage from tensile to shear in intermediate load steps (LS5 and LS6). In this stage, the high probable area is gradually moving toward the mean of the shear class. (c) Dominance of shear cracks at the final load steps (LS7 and LS9). In this stage, the most likely area is concentrated around the mean of the shear class.

Another equally interesting observation is that during the initial loading steps (i.e. LS2 and LS3), the RA components in the shear class are widely distributed between 0 and 60 μs/dB. This is likely due to the fact that microcracks that are generally nucleated during small levels of loads generate AE signals with small amplitudes; this results in larger RA values (see equation (1)). On the other hand, during intermediate and larger loadings (i.e. LS4 onward), the signals are generally of higher amplitudes and consequently RA value varies in a smaller range.

To show the sensitivity of the proposed approach to the sensor location, Figures 11 and 12 show the GMM results for each sensor at LS2 and LS9, respectively. Figure 11 shows that all sensors mostly received tensile signals in LS2, which is consistent with the results shown in Figure 10(a). Figure 12 shows that the overall dominant mechanism in LS9 is shear; however, it can be observed that some sensors (i.e. S1, S2, and S6) indicate a higher probability of tensile mode. These are the sensors closer to the edges and bottom corners of the specimen, where the flexural cracks are dominant as shown in Figure 5. Therefore, sensor-based GMM may estimate the region of flexural and diagonal shear cracks given the sensor location.

Likelihood ratio test

To improve the classification results, a new criterion is proposed to define the cluster boundaries by keeping the data points with higher likelihood into the tensile or shear cluster and associate the rest of data to another cluster, namely, "mixed." With this aim, the likelihood ratio test is used as follows:

$$\Psi = -2\ln\frac{\text{like}\,(\boldsymbol{X})\ \text{for 1st model (Tensile)}}{\text{like}\,(\boldsymbol{X})\ \text{for 2nd model (Shear)}} \rightarrow \begin{cases} \Psi \leqslant \psi_T & \rightarrow \textit{Tensile} \\ \psi_T < \Psi < \psi_S & \rightarrow \textit{Mixed} \\ \Psi \geq \psi_S & \rightarrow \textit{Shear} \end{cases} \quad (12)$$

where Ψ is the likelihood ratio and like$(\boldsymbol{X})$ is the final maximized likelihood in GMM algorithm (i.e., like (X) = The finally maximized $P(X|\lambda_i)$). Now the boundary of clusters can be simply defined by certain values of Ψ, that is, ψ_T and ψ_S for tensile and shear, respectively. In this work, the data points in which the likelihood of being tensile is more than twice the likelihood of being shear are clustered in tensile class. Therefore, $\psi_T = -2\ln(2) = -1.4$ is selected as a boundary of tensile class; likewise, the boundary for shear class is defined by data points in which likelihood of being shear is twice the likelihood of being tensile, i.e. $\psi_S = -2\ln(1/2) = 1.4$. Figure 13 illustrates how the new criterion enhanced the classification by providing an additional option for the operator to select a desired likelihood ratio. The blue diamonds (◇) are associated to the data in which the likelihood of belonging to the tensile mode is twice the likelihood of belonging to the shear mode. The black dots (.) is called "mixed" cluster since it includes the data with less probability of being associated to a specific clusters. The red circles (○) are the data with likelihood higher than twice the likelihood of being tensile.

To track the classification results for all load steps, the number of AE activities associated with each cluster is shown in Figure 14. It can be observed that from load steps 2–4, the majority of AE signals were generated by nucleation of tensile cracks (solid lines in Figure 5). In load steps 5 and 6, the amount of shear cracks began to surpass tensile cracks; this is called transition stage. Finally, a growth of shear cracks was observed in load steps 7–9. Moreover, by comparing these results with the force–displacement hysteresis loops (Figure 4), the following observations can be made: the proposed algorithm shows large AE activity associated to shear mode at the LS7 in which the onset

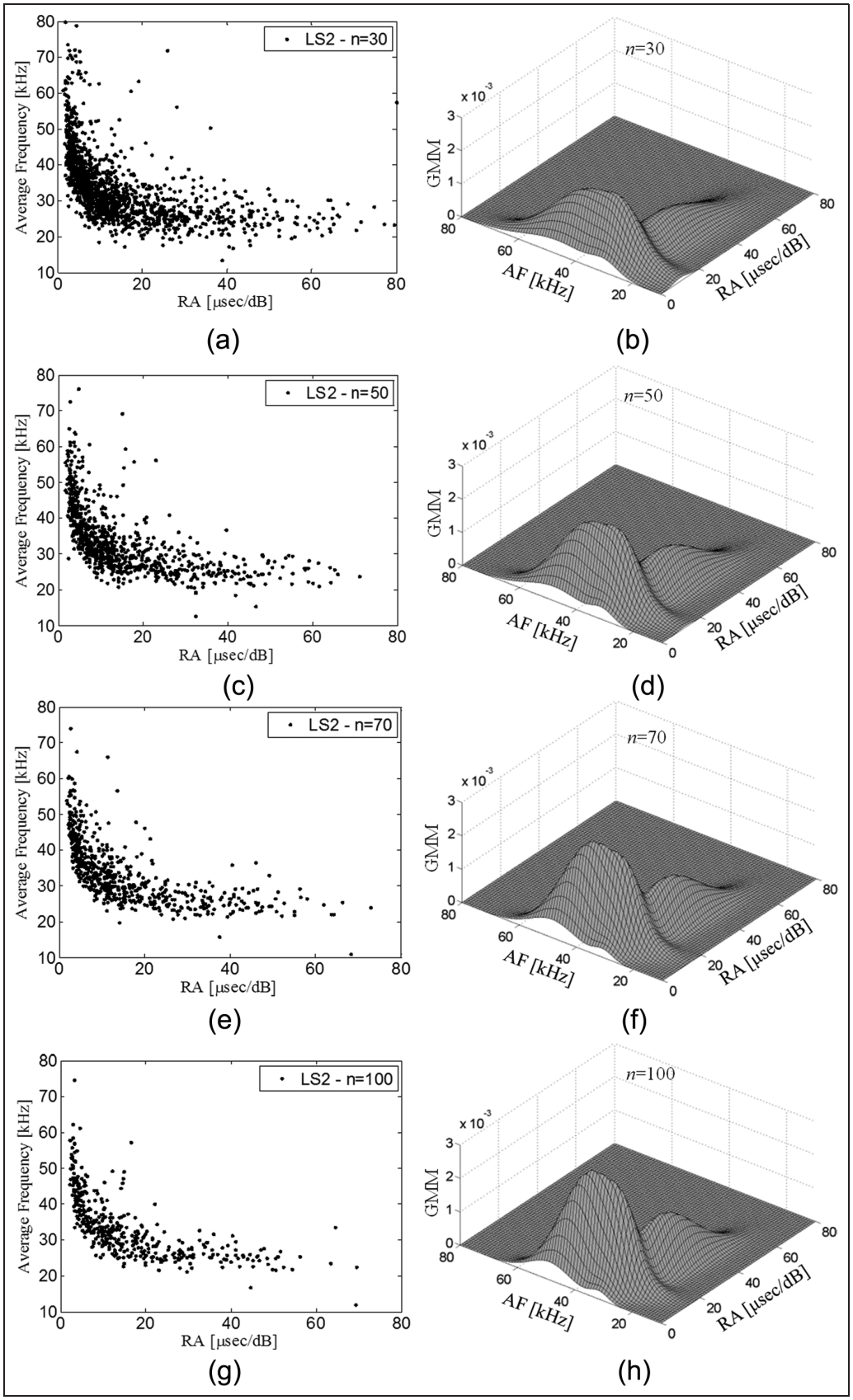

Figure 9. Sensitivity of GMM results to the window span. (a), (b) $\omega_1 = 0.57$, $\omega_2 = 0.43$; (c), (d) $\omega_1 = 0.55$, $\omega_2 = 0.45$; (e), (f) $\omega_1 = 0.58$, $\omega_2 = 0.42$; (g), (h) $\omega_1 = 0.57$ $\omega_2 = 0.43$.
GMM: Gaussian mixture modeling.

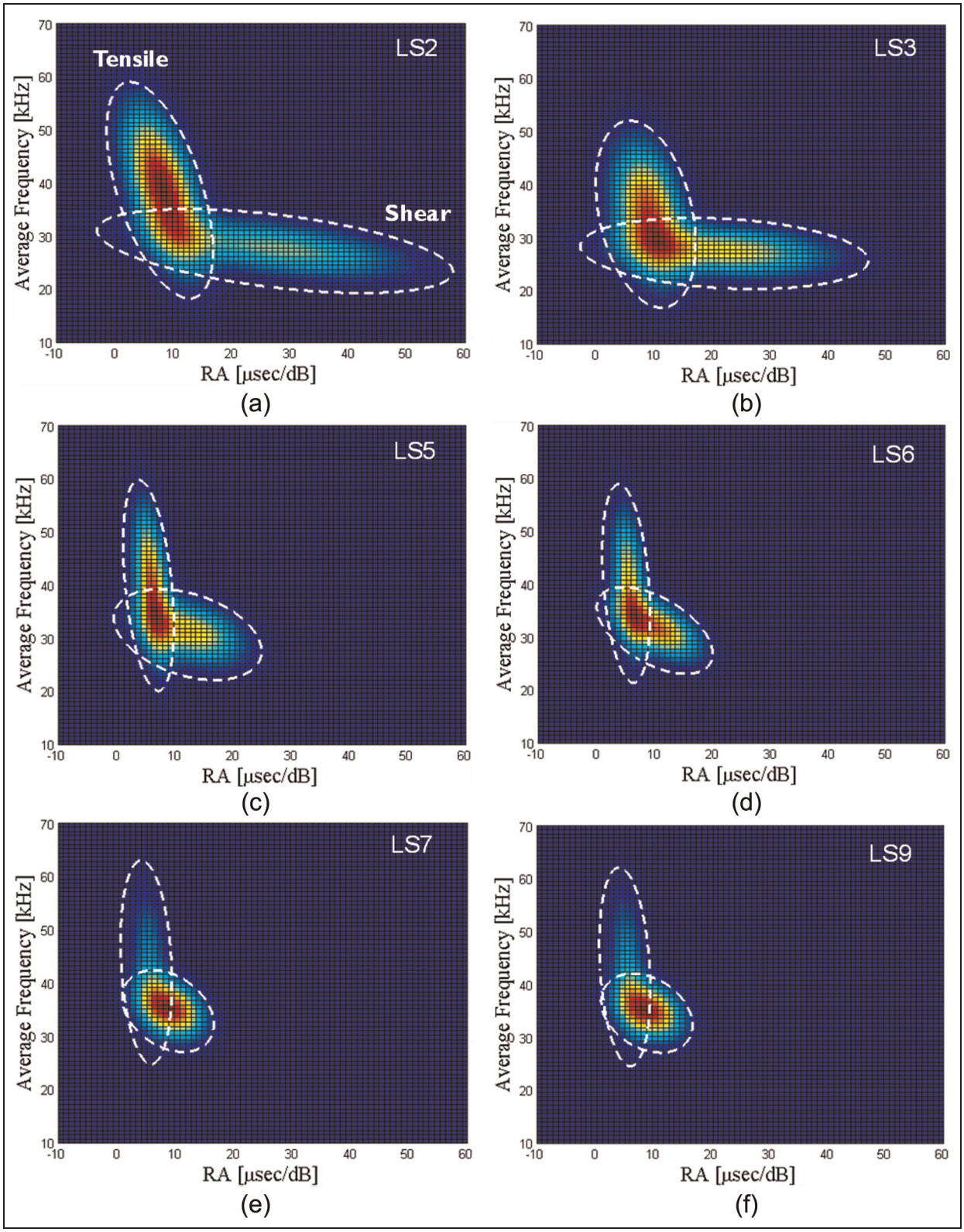

Figure 10. GMM of feature vectors in (a) LS2 and (b) LS3 (dominance of tensile class), (c) LS5 and (d) LS6 (transition stage), and (e) LS7 and (f) LS9 (dominance of shear class).
GMM: Gaussian mixture modeling.

of yielding was observed; then an increase in tensile signals was observed at LS8 due to both nucleation of few additional flexural cracks at the edges of the specimen and mechanism of brittle fracture in compression at the bottom corners (Rocks, 2012); in fact under compressive loads, brittle materials generally fail by the

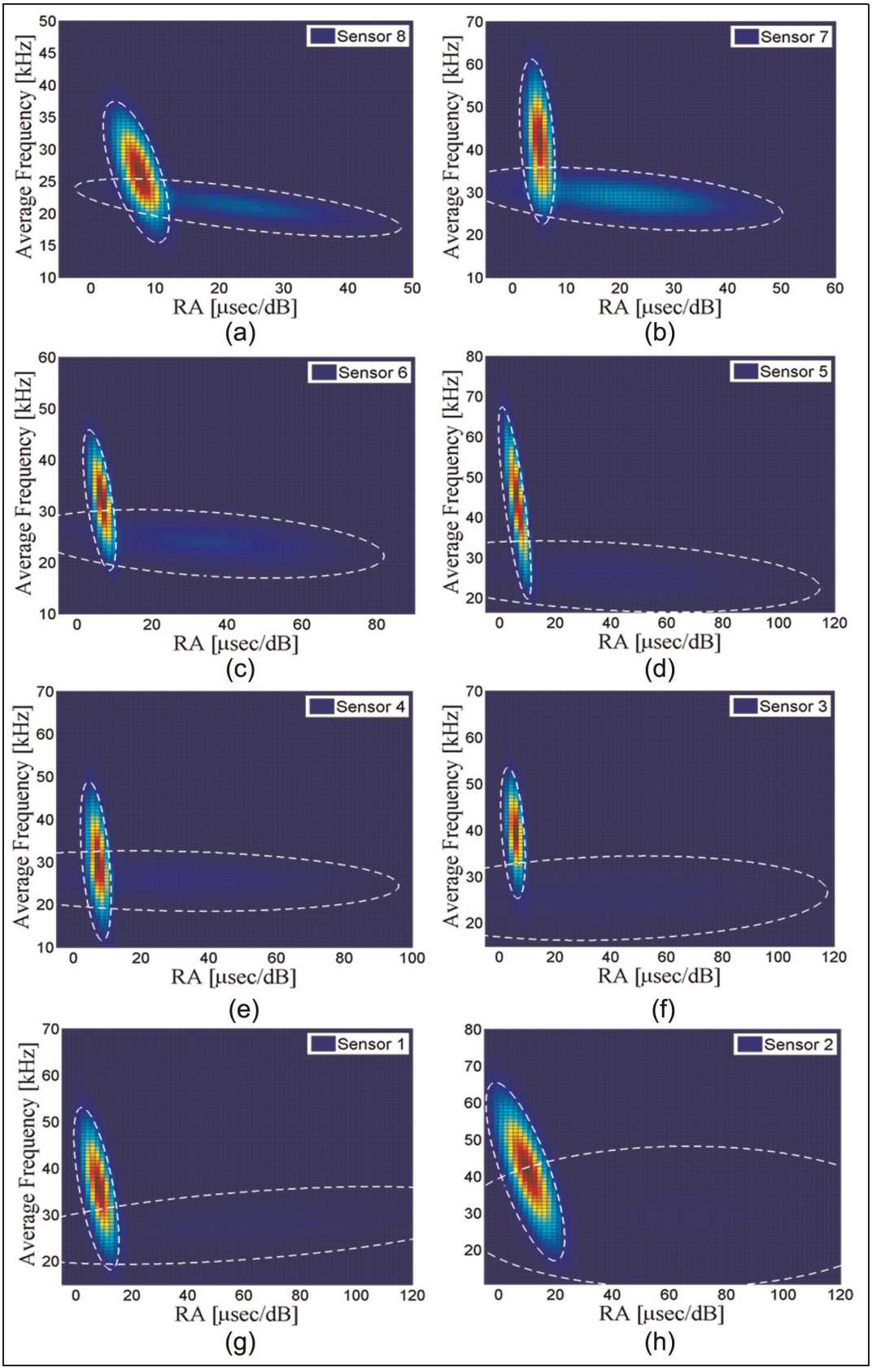

Figure 11. GMM of feature vectors at LS2 for each sensor: (a) Sensor 8, (b) Sensor 7, (c) Sensor 4, (d) Sensor 5, (e) Sensor 6, (f) Sensor 1, (g) Sensor 2, and (h) Sensor 3.
GMM: Gaussian mixture modeling; LS: load step.

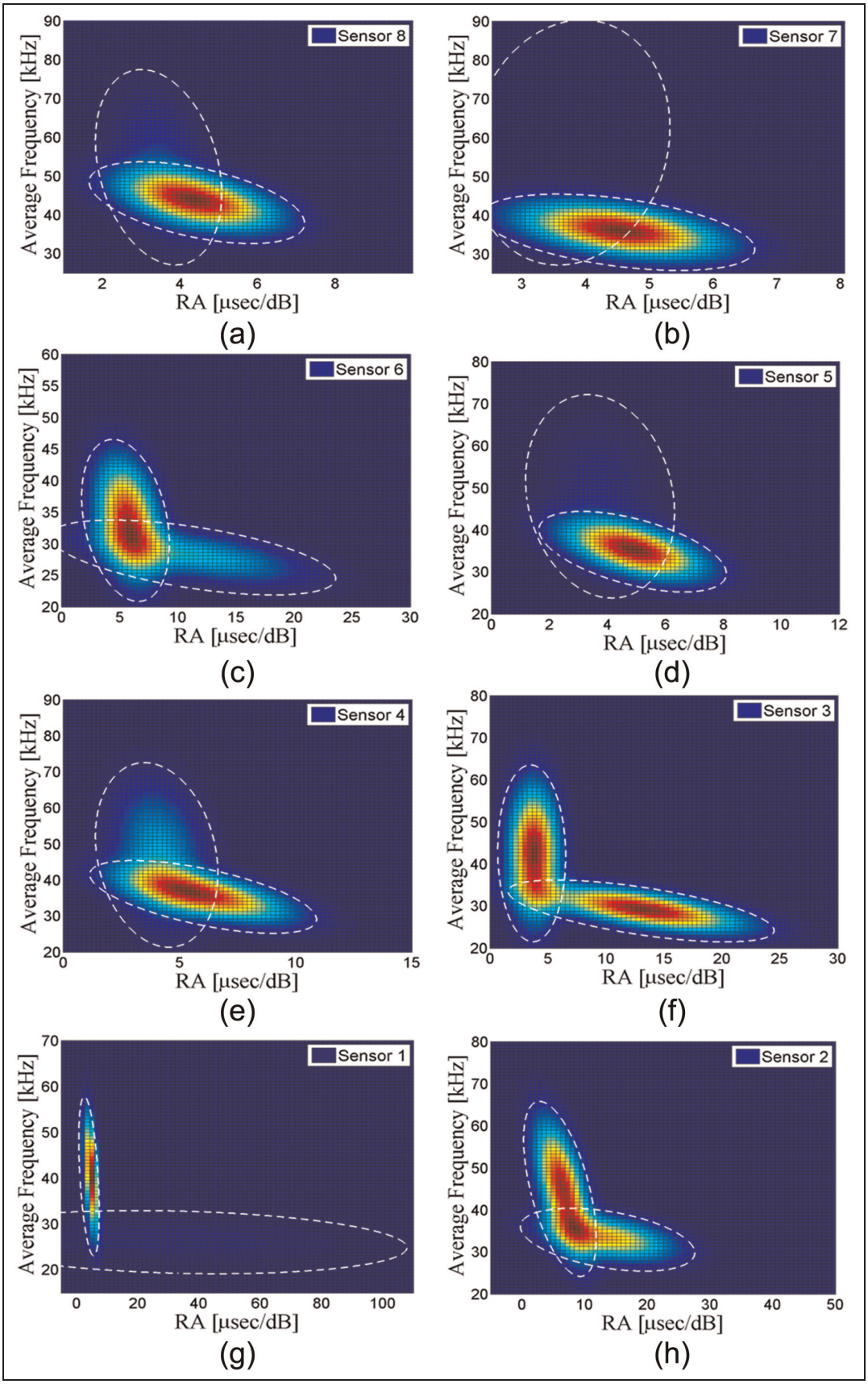

Figure 12. GMM of feature vectors at LS9 for each sensor: (a) Sensor 8, (b) Sensor 7, (c) Sensor 4, (d) Sensor 5, (e) Sensor 6, (f) Sensor 1, (g) Sensor 2, and (h) Sensor 3.
GMM: Gaussian mixture modeling; LS: load step.

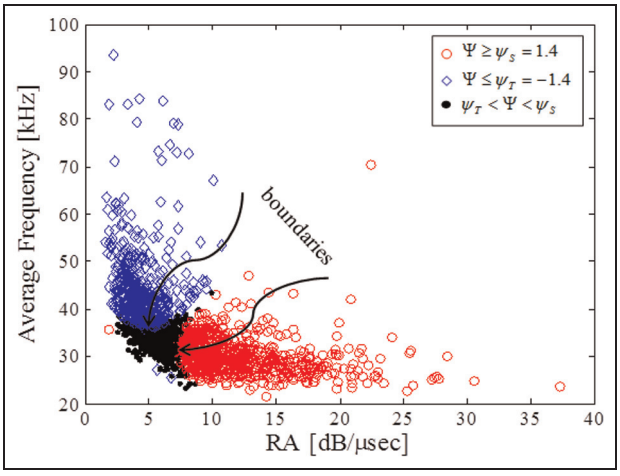

Figure 13. Data selection with higher probability in each cluster in LS6.

Table 1. Percentage of AE activities for each group.

Load step	Drift ratio	Tensile	Shear	Mixed
2	0.11%	59%	33%	8%
3	0.17%	44%	34%	22%
4	0.24%	56%	34%	10%
5	0.32%	40%	38%	22%
6	0.41%	32%	40%	29%
7	0.62%	20%	**74% warning**	6%
8	0.85%	46%	38%	16%
9	1.30%	21%	**63% warning**	16%

AE: acoustic emission.

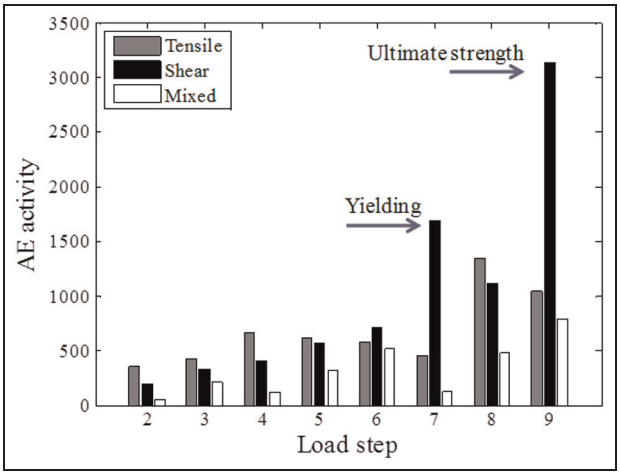

Figure 14. Estimation of crack mode propagation by GMM algorithm.
GMM: Gaussian mixture modeling.

formation of tensile microcracks at microdefects such as cavities and grain interfaces (Nemat-Nasser and Hori, 1993). Finally, at LS9 in which the wall reached its ultimate strength, the GMM indicated significant increases in the population of shear type signals. Table 1 summarizes the percentage of data in each cluster to the total data set in each load step and the associated drift ratio at loading centerline. Based on these observations, a SHM system could be designed for similar structures by defining a threshold on the percentage of shear mode signals. To find this threshold, more specimens should be tested; however, a percentage threshold of 60% for shear class seems reasonable in this study. It highlights yielding and ultimate strength at drift ratios of 0.62% and 1.3%, respectively (bold terms in the table). Therefore, as the percentage exceeds this threshold, an early alarm could be triggered as an indication of significant damage. Thereafter, imminent failure should be prevented by applying an appropriate retrofitting scenario (Bracci et al., 1997; Paterson, 2001; Priestley and Seible, 1996).

Conclusion

In this article, a new approach based on a GMM was presented to perform crack mode classification in RC structures. To validate the proposed system, an experimental study was carried out on a full-scale RC shear wall subjected to a displacement controlled quasi-static reversed cyclic loading. The test specimen was instrumented with eight AE sensors. During testing, two features were extracted from each AE signal detected, that is, RA value and AF. These features were used to populate a data set $\boldsymbol{X}$; then the proposed GMM algorithm was used to partition the data set $\boldsymbol{X}$ in two clusters, tensile and shear. The GMM algorithm was capable to identify three stages during the test: (a) a dominance of tensile (flexural) cracks during the initial load steps, (b) a transition stage during the intermediate load steps, and (c) a dominance of shear cracks during the final load steps. These results were validated by visual inspection. Moreover, a new criterion based on a likelihood ratio test for enhancing crack classification was proposed. Future work should investigate the effects of geometry, specimen ductility, loading rate, and sensor layout.

Funding and Acknowledgments

The authors acknowledge the National Science Foundation (NSF) for providing the financial support under Grant No. CMMI-0829978. The experiments presented herein could not have been completed without contributions from Prof. Whittaker, chair of the department, and the staff of the Structural Engineering and Earthquake Simulation Laboratory (SEESL) of the State University of New York at Buffalo.

Conflict of interest statement

The authors declare that there are no conflicts of interest.

References

ACI Committee 318 (2008) *Building Code Requirements for Structural Concrete and Commentary (ACI 318-08)*. American Concrete Institute (ACI), MI, USA. ISBN: 9780870312649.

Aggelis DG (2011) Classification of cracking mode in concrete by acoustic emission parameters. *Mechanics Research Communications* 38(3): 153–157.

Aggelis DG, Soulioti DV, Sapouridis N, et al. (2011) Acoustic emission characterization of the fracture process in fibre reinforced concrete. *Construction and Building Materials* 25(11): 4126–4131.

Beck P, Lark RJ and Holford KM (2003) Moment tensor analysis of acoustic emission in concrete specimens failed in four-point bending. *Key Engineering Materials* 245–246: 443–450.

Benavent-Climent A, Gallego A and Vico JM (2011) An acoustic emission energy index for damage evaluation of reinforced concrete slabs under seismic loads. *Structural Health Monitoring* 11(1): 69–81.

Bilmes JA (1998) *A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models*. Berkeley, CA: International Computer Science Institute.

Bracci JM, Kunnath SK and Reinhorn AM (1997) Seismic performance and retrofit evaluation of reinforced concrete structures. *Journal of Structural Engineering: ASCE* 123(1): 3–10.

Carpinteri A, Lacidogna G, Niccolini G, et al. (2007) Critical defect size distributions in concrete structures detected by the acoustic emission technique. *Meccanica* 43(3): 349–363.

Colombo S, Forde MC, Main IG, et al. (2005) AE energy analysis on concrete bridge beams. *Materials and Structures* 38(283): 851–856.

Dehghan-Niri E, Farhidzadeh A and Salamone S (2013) Adaptive multisensor data fusion for acoustic emission source localization in noisy environment. *Structural Health Monitoring* 12(1): 59–77.

Dempster AP, Laird NM and Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society Series B: Methodological* 39(1): 1–38.

Easton VJ and McColl JH (1997) *Statistics glossary v1.1*. The STEPS Project, STatistical Education through Problem Solving, University of Glasgow, UK.

ElBatanouny MK, Larosche A, Mazzoleni P, et al. (2012) Identification of cracking mechanisms in scaled FRP reinforced concrete beams using acoustic emission. *Experimental Mechanics*. DOI: 10.1007/s11340-012-9692-3.

Farhidzadeh A and Salamone S (2012) Introducing sifted b-value analysis and a new crack classification for monitoring reinforced concrete shear walls by acoustic emission. In: *54th acoustic emission working group meeting*, Princeton, NJ, USA, May 20-23, 2012.

Farhidzadeh A, Dehghan-Niri E, Salamone S, et al. (2012a) Monitoring crack propagation in reinforced concrete shear walls by acoustic emission. *Journal of Structural Engineering: ASCE*. DOI: 10.1061/(ASCE)ST.1943-541X.0000781.

Farhidzadeh A, Salamone S, Dehghan-Niri E, et al. (2012b) Damage assessment of reinforced concrete shear walls by acoustic emission. In: *NDE/NDT for highways and bridges: structural materials technology (SMT)*, NY, USA, August 21–24, 2012, pp. 74–81.

Farhidzadeh A, Salamone S, Luna B, et al. (2013) Acoustic emission monitoring of a reinforced concrete shear wall by b-value based outlier analysis. *Structural Health Monitoring* 12(1): 3–13.

Grosse C, Reinhardt H and Dahm T (1997) Localization and classification of fracture types in concrete with quantitative acoustic emission measurement techniques. *NDT & E International* 30(4): 223–230.

Grosse CU and Ohtsu M (2008) *Acoustic Emission Testing—Basics for Research-Applications in Civil Engineering*. Berlin & Heidelberg: Springer-Verlag.

Hastie T, Tibshirani R and Friedman J (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Stanford, CA: Springer, pp. 272–279.

Jahanshahi MR, Masri S, Padgett CW, et al. (2013) An innovative methodology for detection and quantification of cracks through incorporation of depth perception. *Machine Vision and Applications* 24: 227–241.

JCMS-IIIB5706 (2003) *Monitoring Method for Active Cracks in Concrete by Acoustic Emission*. Federation of Construction Material Industries, Japan, pp. 23–28.

Kurz JH, Finck F, Grosse CU, et al. (2006) Stress drop and stress redistribution in concrete quantified over time by the b-value analysis. *Structural Health Monitoring* 5(1): 69–81.

Luna B (2013) *Seismic response of low aspect ratio reinforced concrete walls for building and safety-related nuclear applications*. PhD Thesis, University of Buffalo, Buffalo, NY.

Miller RK, Hill E and Moore PO (2005) *Nondestructive Testing Handbook Acoustic Emission Testing*. 3rd ed., vol. 6. Columbus, OH: American Society for Nondestructive Testing, pp. 1–25.

Nemat-Nasser S and Hori M (1993) *Micromechanics: Overall Properties of Heterogeneous Materials*. Amsterdam: North-Holland.

Ohno K and Ohtsu M (2010) Crack classification in concrete based on acoustic emission. *Construction and Building Materials* 24(12): 2339–2346.

Ohtsu M (2010) Recommendation of RILEM TC 212-ACD: acoustic emission and related NDE techniques for crack detection and damage evaluation in concrete. *Materials and Structures* 43(9): 1187–1189. DOI: 10.1617/s11527-010-9640-6.

Paterson J (2001) *Seismic retrofit of reinforced concrete shear walls*. Master's Thesis, McGill University, Montreal, QC, Canada.

Permuter H, Francos J and Jermyn I (2006) A study of Gaussian mixture models of color and texture features for image classification and segmentation. *Pattern Recognition* 39(4): 695–706.

Povinelli RJ, Johnson MT, Lindgren AC, et al. (2004) Time series classification using Gaussian mixture models of reconstructed phase spaces. *IEEE Transactions on Knowledge and Data Engineering* 16(6): 779–783.

Priestley MJN and Seible F (1996) Design of seismic retrofit measures for concrete and masonry structures. *Construction and Building Materials* 9(6): 365–377.

Rabiner LR and Juang BH (1986) An introduction to hidden Markov models. *IEEE ASSP Magazine*, January, 4–16.

Rammohan R and Taha MR (2005) Exploratory investigations for intelligent damage prognosis using hidden

Markov models. In: *2005 IEEE international conference on systems, man and cybernetics*, Albuquerque, NM, 12–12 October, vol. 2, pp. 1524–1529.

Reynolds DA and Rose RC (1995) Robust text-independent speaker identification using Gaussian mixture speaker models. *IEEE Transactions on Speech and Audio Processing* 3(1): 72–83.

Reynolds DA, Quatieri TF and Dunn RB (2000) Speaker verification using adapted Gaussian mixture models. *Digital Signal Processing* 10(1–3): 19–41.

Richiardi J and Drygajlo A (2003) Gaussian mixture models for on-line signature verification. In: *Proceedings of the 2003 ACM SIGMM workshop on Biometrics methods and applications*—WBMA, Berkeley, CA, 8 November, pp. 115–122.

Rocks JF (2012) *Large scale testing of low aspect ratio reinforced concrete walls*. MSc Thesis, Department of Civil, Structural and Environmental Engineering, University at Buffalo, Buffalo, NY.

Salamone S, Bartoli I, Phillips R, et al. (2011) Health monitoring of prestressing tendons in post-tensioned concrete bridges. *Journal of the Transportation Research Board* 2220: 21–27.

Salamone S, Veletzos MJ, Lanza di Scalea F, et al. (2012) Detection of initial yield and onset of failure in bonded post-tensioned concrete beams. *Journal of Bridge Engineering* 17: 966–974.

Shigeishi M (2001) Acoustic emission moment tensor analysis: development for crack identification in concrete materials. *Construction and Building Materials* 15(5–6): 311–319.

Shiotani T, Yuyama S, Li ZW, et al. (2001) Application of AE improved b-value to quantitative evaluation of fracture process in concrete materials. *Journal of Acoustic Emission* 19: 118–133.

Soulioti D, Barkoula NM, Paipetis A, et al. (2009) Acoustic emission behavior of steel fibre reinforced concrete under bending. *Construction and Building Materials* 23(12): 3532–3536.

Stauffer C and Grimson WEL (1999) Adaptive background mixture models for real-time tracking. In: *Proceedings of the 1999 IEEE computer society conference on computer vision and pattern recognition*. Fort Collins, CO, 23-25 June, 1999, pp. 246–252. IEEE Computer Society.

Terejanu G, Singla P, Singh T, et al. (2008) Uncertainty propagation for nonlinear dynamic systems using Gaussian mixture models. *Journal of Guidance Control and Dynamics* 31(6): 1623–1633.

Vanniamparambil PA, Bartoli I, Hazeli K, et al. (2012) An integrated structural health monitoring approach for crack growth monitoring. *Journal of Intelligent Material Systems and Structures* 23(14): 1563–1573.

Yuyama S, Li Z, Ito Y, et al. (1999) Quantitative analysis of fracture process in RC column foundation by moment tensor analysis of acoustic emission. *Construction and Building Materials* 13: 87–97.

Zhou W, Kovvali N, Reynolds W, et al. (2009) On the use of hidden Markov modeling and time-frequency features for damage classification in composite structures. *Journal of Intelligent Material Systems and Structures* 20(11): 1271–1288.